

Are Infinite Explanations Self-Explanatory?

Alexandre Billon

This is the pre-peer-reviewed version of a paper forthcoming in *Erkenntnis*
please quote the published version

May 31, 2021

Abstract

Consider an infinite series whose items are each explained by their immediate successor. Does such an infinite explanation explain the whole series or does it leave something to be explained? Hume arguably claimed that it does fully explain the whole series. Leibniz, however, designed a very telling objection against this claim, an objection involving an infinite series of book copies. In this paper, I argue that the Humean claim can, in certain cases, be saved from the Leibnizian “infinite book copies” objection, and that this provides an interesting way to defuse some cosmological arguments for the existence of God and to give a non-theistic but complete explanation of the Universe. In the course of my argumentation, I also show that circular explanations can be “self-explanatory” as well: explaining two items *by each other* can explain the couple of items tout court.

1 Introduction

Consider a series $(u_i)_{i \in I}$ whose items are each (fully) explained by their immediate successor. Items can be states of the Universe, events, propositions, contingent beings, instances of coming to existence, etc. I can be the set of non-null natural numbers \mathbb{N}^* , in which case $(u_i)_{i \in I}$ constitutes an *infinite chain of explanations* or, for short, an *infinite explanation*.¹ I can also be the the ring of integers modulo n , $\mathbb{Z}/n\mathbb{Z}$,² in which case $(u_i)_{i \in I}$ constitutes a *circular chain of explanations* or, for short, a *circular explanation*. Question: Is it the case that by explaining the transitions from each item to its antecedent, we ipso facto explain all the items of the series? If it is, and only if it is, we will say that the series is

¹As an item is explained by its successor, the chain is, so to speak, descending. Infinite series in which each term explains the next can be called “ascending infinite explanations”. They raise interesting, sui-generis problems (see for example Pruss (2018b, III)) that cannot be tackled within the scope of this paper.

²The ring of integers modulo n ($\mathbb{Z}/n\mathbb{Z}$) is the n first integers $0, 1, \dots, n-1$ endowed with the addition and multiplication operations, and where (to put it rather roughly) it is assumed that for all x , $x = x + n$.

“self-explanatory”.³

The issue is important because, among other things, a significant kind of cosmological arguments rely on the premise that infinite and circular explanations are not self-explanatory. The kind of arguments I am thinking of come in great variety of guises, some of which have been finely wrought to defuse subtle objections. For the present purposes, a rough and simple version will suffice. Let us call “self-explanatory” an item that literally explains itself (notice that the term “self-explanatory” does not mean exactly the same thing when applied to a series of explanations and to an item in the series). The rough and simple version of the cosmological argument relies on the three following premises:

- (1) Everything that is not self-explanatory is explained by something else.⁴
- (2) Infinite explanations are not self-explanatory (by explaining each item of an infinite series by its successor, one does not explain the whole series of items).
- (3) Circular explanations are not self-explanatory (by explaining each item of a circular series by its successor, one does not explain the whole series of items).

Now say that an explanation is *complete* if it leaves nothing to be explained. By a well-worn reasoning, (1) entails that a complete explanation of a given, initial item can take just three forms: it will either be (i) a self-explanatory *infinite* explanation (ii) or a self-explanatory *circular* explanation (iii) or a finite explanation that ends with a *self-explanatory item*. (2) entails that option (i) will not work. (3) that option (ii) will not work either. Which only leaves the option (iii). Accordingly, if there is a complete explanation of our initial item, there must be a self-explanatory item — a supposed privilege of God and other godly things (see Rowe (1970); Pruss (1998)).⁵ It is common to distinguish causal from metaphysical or “grounding” explanations (see for example Fine (2012)). As I understand it, this version of the cosmological argument admits both interesting grounding and causal readings and we will refer to both readings in what follows.⁶

³We might want to say that a circular or an infinite explanation that is not self-explanatory is vicious, and involve either a *vicious circle* or a *vicious regress*. As emphasised by Bliss (2013), the vicious character of an explanation depends, however, on the role we assign to this explanation. Importantly, a non-self-explanatory circular or an infinite explanation will be vicious if it purports to explain the whole series. It is, however, not clear that it will be vicious otherwise. We shall see in fn.22 that the main conclusion of this paper can in fact help one defuse some recent influential arguments for the generally vicious character of infinite explanations.

⁴This premise is a simple version of the “Principle of Sufficient Reason”. It would be denied by someone who claims that there are brute facts, facts that are not explained by anything.

⁵The existence of a self-explanatory item, as that of a circular chain of (transitive) explanations entails that there are things that literally explain themselves and that explanation is not irreflexive. This claim has been disputed (more or less) independently of the cosmological argument (see e.g. Oppy (2006, 277-8) and Guigon (2015) for arguments for irreflexivity and Hudson (1997), and Pruss (2006, VI) for arguments against it). I could not rehearse these rich debates without far outreaching the scope of the present paper.

⁶This argument is inspired by Leibniz (1989)’s and Clarke (1998)’s arguments, but Leibniz and Clarke did not have in mind causal or purely causal explanations. They were interested in the first reasons rather than in the first causes. Simplifying somewhat, I will assume that these reasons are metaphysical grounds

The issue has also been quite disputed in this context. Hume for example claimed that there is nothing left to be explained once the items composing a whole are all explained (Hume, 1907, IX), which seems to imply that infinite explanations are self-explanatory.⁷

Yet Leibniz famously argued that infinite explanations are not self-explanatory. Leibniz (1989) indeed provided a nice counterexample to the thesis that infinite explanations are self-explanatory, a counterexample that we can summarise as follows:

The Regular Scribes. Consider an infinite series of book copies. Book copy #1 is a copy of the Elements of Geometry, its content is explained by the fact that it was copied before by a *regular scribe* from an older book copy #2; this older book copy #2 is a copy of the Elements of Geometry, its content is explained by the fact that it was copied, before, by a *regular scribe* from a yet older book copy #3, on so on to infinity... where a regular scribe is someone who makes a faithful copy of the book he is given.

Indeed, by explaining the transition from one book copy of the series to its antecedent (i.e. by mentioning that scribes take a book copy and make another, antecedent one with the same content) one explains why all book copies have the same content. One does not, however, ipso facto explain why they are all copies of the elements geometry rather than (say) the Bible or why this content existed in the first place.⁸ One does not explain the whole series.

Leibniz's counterexample — call it Leibniz's "Regular Scribes" case — probably goes some way toward explaining the contemporary quasi-consensus. Indeed, the claim that infinite explanations are not self-explanatory has become orthodox, both among contemporary theists and atheists. In a recent essay that surveys the classical purportedly complete

but not causes.

There are ways to resist the conclusion that do not involve rejecting (2) or (3). One might, for example, deny that self-explanatory items need to be divine (as Dasgupta (2016) does, in the case of metaphysical explanations) or deny that anything — even godly things — can be self-explanatory (as Russell (1927) does in the case of causal explanations).

There are also other kinds of cosmological arguments, most notably the Kalām argument and Aquinas's third way. Both rely on the claim that an infinite explanation is impossible or in some way *vicious*.

⁷The whole, Hume said, is "sufficiently explained in explaining the cause of the parts." The principle, to the effect that by explaining the parts of a whole, one explains the whole itself, is sometimes called the "Hume-Edwards principle" (the name comes from Rowe (1970)). In fairness, it is not totally clear to me that Hume really endorsed this principle. By claiming that the whole is "sufficiently explained," Hume might have wanted to convey the idea that that it is not even apt for being explained, or, even more plausibly in my view, the idea that it is brute (i.e. unexplained and apt for being explained) *but that we should accept it as brute* (see Dasgupta (2016) on the distinction between things that are brute and those that are not even apt for being explained).

⁸As Leibniz (1989) puts it, our explanation will never be "complete (...) for we can always ask: Why have there always been such books? Why were these books written? Why were they written in the way they were?"

explanations of the Universe, Derek Parfit neatly summarises the common wisdom:⁹

Some believe that, if all events were caused by earlier events, everything would be explained. That, however, is not so. Even an infinite series of events cannot explain itself. We could ask why this series occurred, rather than some other series, or no series (Parfit (2011, vol.II, 623), cf. also, p. 644 for the same claim about non-causal explanations)

Leibniz’s counterexample does show that *some* infinite explanations are not self-explanatory. However, and despite Leibniz’s contention, it does not show that infinite explanations are *never* self-explanatory. In this paper, I will argue that in some cases they are in fact self-explanatory, and that such cases might be relevant to the actual world, allowing one to provide complete non-theistic explanations that are at least as plausible as theistic complete explanations.

Here is how I will proceed. In section 2, I will put forward a few simple examples of self-explanatory infinite explanations, most notably the Monomaniacal Scribes case and the Stick-Adjusters case. I will then consider an objection from Alexander Pruss against the possibility of such explanations, an objection which relies on the premise that circular explanations are never self-explanatory. I argue against this premise, relying on circular versions of the Monomaniacal Scribes case and the Stick-Adjusters case (section 3). In section 4, I will put forward general necessary and sufficient conditions for an infinite (or indeed a circular) explanation to be self-explanatory. The necessary conditions show that such explanations will always be quite peculiar, displaying a form of “historical irrelevance”. In the last sections, I argue that their peculiarity should not, however, have us consider them as defective or illegitimate (section 5), and that such explanations indeed provide a plausible alternative to theistic explanations, and are accordingly relevant to the discussion of the cosmological argument (section 7). I also answer an important objection concerning the link between explanation and necessitation (section 6).

2 Some Self-Explanatory Infinite Explanations

Why does Leibniz book copies case show that some infinite explanations are not self-explanatory? Why do we want to say that one does not explain the whole series by explaining the transitions from the book content of each book from that of its antecedent? It is because what explains these transitions — namely the specified behavior of the regular scribes who take any book and make a book copy with the same content — *does not (fully) explain the whole series of book contents*. And why do we want to say that it doesn’t (fully) explain it? It is, I take it, because it does not necessitate it: we could have an infinite series of books such that each book has been copied by a regular scribe from its successor but

⁹See Rowe (1970), Rowe (2001, 24-5), Pruss (1998, 2018a,b), Pruss (2006, III) and Swinburne (2004, 139-43) Cameron (2008, 12) for other influential endorsements of the view.

such that not all books are copies of the the Elements of Geometry instead (they could for example all be copies of Bible instead).

In this section and the next one, I will assume that it is *only because* what explains these transitions fails to necessitate the whole series that this infinite explanation is not self-explanatory, and connectedly that *an explanation of the same form would be self-explanatory if what explains the transitions necessitated the whole series* — call that the necessitation condition. I will try to remain neutral on the kind of necessity at stake here (is it epistemic, metaphysical or yet of another kind?), which might depend both on the kind of explanation we consider and on our account of that kind of explanation.

We will question this necessitation condition later (in section 6). Once this necessitation condition is granted, however, it is not difficult to come up with an example of a self-explanatory infinite explanation. Consider the following variation on Leibniz’s book copy case:

The Monomaniacal Scribes. We have an infinite series of book copies. Book copy #1 is a copy of the Elements of Geometry because it was copied before by a *monomaniacal scribe* from an older book copy #2; the older book copy #2 is a copy of the Elements of Geometry because it was copied, before by a *monomaniacal scribe* from a yet older book copy #3, on so on to infinity... where a monomaniacal scribe is someone who, instead of making a faithful copy of the book he is given, always creates a copy of the Elements of Geometry, whatever book copy he is given.

This explanation, I take it, does seem self-explanatory. By explaining each transition one seems to explain why each book has the content it has. And unlike in Leibniz’s Regular Scribes case, moreover, it seems that what explains these transitions, namely the behavior of Monomaniacal Scribes, do necessitate each item of the series. Necessarily, if all the scribes behave as specified, all books must be copies of the Elements of Geometry.¹⁰

This example, one might contend, is somehow trivial or degenerate, however. In particular, the content of each book does not really depend on the content of the subsequent book in the following, “difference-making” sense: changing the latter would not make any difference to the former. Connectedly, it seems that the infinite explanation is somehow idle here. One could explain the first item by the action of the first scribe without mentioning either the second book or the subsequent scribes.

Here is another, non-degenerate example, an example that involves the successive lengths of a plastic stick rather than the contents of books and in which the items do depend on

¹⁰One might wonder whether the behavior of the scribes necessitates (and indeed explains) why all the books exist and have a content at all. It does, provided that we make it explicit that our monomaniacal scribes create a copy of the Elements of Geometry whatever book copy they are given and even, in particular, if they are given a book with an empty content or no book at all.

each other. In order to make it simpler, and although this supposition is not necessary (see fn.11), we will suppose that possible plastic sticks are bounded in size by an upper bound B and a lower bound b .

The Stick Adjusters. Consider the infinite series of lengths $(l_i)_i$ of a given plastic stick (the length of plastic sticks is necessarily between b and B). Length l_1 is l , which is explained by the fact that the stick has been adjusted before by a stick-adjuster from a former state in which it had length l_2 , which length is explained by the fact that it has been adjusted earlier by another stick-adjuster from state in which it has length l_3 , and so on to infinity... where a stick-adjuster is someone who takes a stick of length x and

- (i) if $l \leq x \leq B$, cuts it by $\frac{x-l}{2}$ so that its size becomes $x - \frac{x-l}{2}$.
- (ii) if $b \leq x \leq l$, extends it by $\frac{l-x}{2}$ so that its size becomes $x + \frac{l-x}{2}$.

It takes little reflection to realize that necessarily, if the stick-adjusters behave as specified, the stick will always be exactly l long. Indeed, it can be checked that $l_{i+j} = l + 2^j(l_i - l)$, so if for some i , $l_i > l$, the series $(l_{i+j})_{j \geq 1}$ would have no upper bound, which is absurd by construction (unless otherwise mentioned indexes are natural numbers greater than or equal to one). Similarly, if for some i , $l_i < l$, the series $(l_{i+j})_{j \geq 1}$ would have no lower bound, which is also absurd by construction.¹¹ Just like in the preceding Monomaniacal Scribe case, this infinite explanation thus satisfies our necessitation condition: what explains the transitions from one length to its antecedent necessitates all lengths of the stick. It makes it necessary that the stick is always l -long, and it also seems, intuitively, to explain everything there is to explain. This infinite explanation is thus self-explanatory.¹²

This infinite explanation is not degenerate, however, as the length of the stick at one step really depends on that of the stick at the next step: changing the length of the stick at step $\#i + 1$ would indeed change the length of the stick at the preceding step

¹¹We could dispense with the supposition that the size of sticks is generally bounded by stipulating that there are real numbers b and B such that if a stick-adjuster is given a stick whose length is bigger than B (res. smaller than b) he cuts it to make a B -long stick (res. he extends it to make a b -long stick). The reader can check that if for some i , $l_i \neq l$, there would be some j such that $l_{i+j} > B$ or $l_{i+j} < b$, which is absurd because there is no length k such that a stick-adjuster can make a stick of a length $< b$ or $> B$ out of a stick of length k .

¹²One might quarrel with the claim that our necessitation condition is satisfied in the Stick-Adjusters case on the ground that a form of nonstandard analysis such as Robinson (1961)'s could be true, and that if it were true, then the stick could either be l -long at each step, or infinitesimally smaller than l at each step, or again, infinitesimally bigger than l at each step. In response, we might weaken all our claims and say that they only hold approximately (the behavior of the stick-stretchers necessitates and explains that the length of the stick is always approximately l , that is, l modulo an infinitesimal ϵ), which, I believe, would still lead to significant conclusions. We could also stipulate that stick-adjusters make sure that all lengths are real numbers (and hence obey classical, rather than nonstandard, analysis). In any case, this objection does not apply to the Wheel-Turners case below.

i . Similarly, one cannot fully explain the first length l_1 of the stick by the action of the first Stick-Adjuster (without mentioning the second length of the stick or the actions of all subsequent Stick-Adjusters).

We can come up with other, maybe simple examples of non-degenerate self-explanatory infinite explanations. Consider the Wheel-Turners.

The Wheel-Turners. Take a wheel that is divided in four identical numbered sectors (respectively 0, 1, 2, 3). The sector s_1 on which the wheel has just landed is 1 because the wheel has just been turned by a wheel-turner from a former sector s_2 , and it was in that sector because the wheel had been turned by a wheel-turner from a former sector s_3 ... where a wheel-turner is someone who takes a wheel that has landed on sector x ($x \in \mathbb{Z}/4\mathbb{Z}$) and turns it so that it lands on sector $f(x) = 2x - 1$ ($f(x) \in \mathbb{Z}/4\mathbb{Z}$).

As $f(0) = f(2) = 3$, $f(3) = 1$ and $f(1) = 1$,

- if the wheel lands on 1 it will always stay on 1 when it is turned again,
- and the wheel will always land on one (whatever the starting point) provided that it has been turned at least twice.

So for all i , $s_i = 1$. This, moreover, holds necessarily provided that the wheel-turners act as specified. Accordingly, the necessitation condition is satisfied. This explanation seems self-explanatory as well.

The three examples we have put forward so far involve *constant* items. In the Monomaniacal Scribes case, all books have the same content. In the Stick-Adjusters all sticks have the same length and in the Wheel-Turners cases, the wheel remains on the same section. It is important to notice that not all self-explanatory infinite series are likewise constant, however. Consider the following variations on the cases above.

The Sliding Monomaniacal Scribes. Consider a case that is just like the ordinary Monomaniacal Scribes case above except the scribes are sliding monomaniacal scribes, where such a sliding monomaniacal scribe # i creates a copy of the Elements of Geometry and erases its i last words, so that the book copy # $i + 1$ is a version of the Elements of Geometry amputated of i words.

The reader can check that this is a (degenerate) case of non-constant self-explanatory infinite explanation.

The Sliding Stick Adjusters. Consider case that is just like the ordinary Stick-Stretches case above except that stick adjusters are sliding stick-adjuster, where a sliding stick-adjuster s_i someone who, after having adjusted the stick, takes away whatever is glued to the stick and glues a i -long ribbon to one of its extremities instead.

If we take the items in this case to be the length of stick plus that of the ribbon that was glued to it, we have a case of non-degenerate self-explanatory infinite explanation whose items are non-constant.

While in the Regular Scribes example, the infinite explanation explains very little of the whole series (by explaining the transitions, one explains why all the books of the series have the same content but one does not explain anything else), in the five examples above the infinite explanation fully explains the whole series. It explains everything that needs to be explained. There are intermediate cases, though, in which a lot, but not everything, seems to be explained. Consider the following Recalcitrant Scribe example:

The Recalcitrant Scribes. Consider a case that is similar to the other Scribes cases except that the scribes are *recalcitrant scribes*, where a recalcitrant scribe is one who, when he is given a copy of the Bible (res. the Elements of Geometry) creates a copy of the Elements of Geometry (res. the Bible).

By explaining the transitions, here, one explains why we have a cyclical series in which copies of the Elements of Geometry succeed to copies of the Bible. However, one does not ipso facto explain why the book copy #1 is the Elements of Geometry rather than the Bible. We might want to say that this infinite series is partially but not fully self-explanatory, or that it is self-explanatory within a period 2 (it explains why book copy #1 or its successor is the Elements of Geometry).¹³

3 Infinite and Circular Explanations

Most philosophers who argue against self-explanatory infinite explanations ground their claim, like Leibniz, on some counterexamples.¹⁴ Their arguments thus rely on inductive or abductive generalizations. Our examples of self-explanatory infinite explanations suggest that these arguments fail. Some forms of infinite explanations are not self-explanatory, but others are.

Pruss (1998, 156) has, however, put forward an interesting *deductive* argument against self-explanatory infinite explanations. If cogent, his argument would show that there is something wrong with our (purported) counterexamples (see also Pruss (2018b, II) and Pruss (2006, III)). Pruss (1998) does start by providing a couple of counterexamples to the claim that all infinite explanations are non-self-explanatory (what he calls the “cannon-ball” case and the “chickens and eggs” example), and I agree, even if not quite for the same reasons, that his counterexamples succeed. The deductive argument he adds is complex, but it will not be necessary to rehearse it in all its details. It rests on the close connection

¹³This example is similar to Pruss (1998, 154)’s chickens and eggs example. My argument for the non-self-explanatory character of the series is, however, different.

¹⁴Beside Leibniz, Rowe (1970); Pruss (1998) have for example put forward counterexamples.

between infinite and circular explanation, and on the premise that circular explanations of one item by another and vice-versa are never self-explanatory (by explaining one item by the other and vice-versa one does not fully explain the two items taken together). I disagree with this premise just like I disagree with the claim that all infinite explanations are not self-explanatory, and for very similar reasons. Again, it is true that *some* circular explanations are not self-explanatory.

Regular Scribes (circular case). Consider two book copies. Book copy #1 is a copy of the Elements of Geometry, its content is explained by the fact that it has been copied in 1999, by a regular scribe, from a book copy #2. Book copy #2 is also a copy of the Elements of Geometry, and its content is explained by the fact that it has been copied yesterday from book copy #1 by a regular scribe who then traveled through time to 1999 with book copy #2.

This circular explanation is not self-explanatory because the behavior of the scribes explains why both books have the same content, not why this content is that of the Elements of Geometry rather than, say, that of the Bible. And it fails to explain that because, and, arguably, only because, it does not make it necessary that both books are copies of the Elements of Geometry. But not all circular explanations fail to be self-explanatory. In the example above, if, we replace our regular scribes by what we have called “monomaniacal scribes”, who always make copies of the Elements of Geometry, whatever they take as input, then it seems that the behavior of the scribes will make it necessary that book copies #1 and #2 are both copies of the Elements of Geometry and that it will explain the content of both books. This circular explanation will accordingly be self-explanatory.¹⁵

If we assume again that circular explanations of the same form as the Regular Scribes would be self-explanatory if only what explains the transitions could necessitate all items, the circular version of the Stick-Adjusters case seems likewise self-explanatory:

The Stick-Adjusters (circular case). Consider the length l_1 and l_2 of a stick in 2020 and 1999 respectively. The length l_1 is l , which is explained by the fact that the stick has been adjusted at the end of 1999 by a stick-adjuster and that in 1999 it had length l_2 , which length is in turn explained by the fact

¹⁵Most of the so-called paradoxes of time travel (I travel to the past and kill my grandfather and thus could not exist) and “hypodoxes” of time travel (I travel to the past with a unique recipe my grandfather gave me, I leave it on a table where my grandfather discovers it for the first time and copies it; there is no inconsistency here but the existence and content of the recipe seems to be left totally unexplained, see Eldridge-Smith (2007)) rest on such non-self-explanatory circular explanations. Hypodoxes of time-travel are in fact nothing but consistent scenarios of time-travel involving a circular explanation that is not self-explanatory (and that accordingly seems to leave something unexplained and “indeterminate”). I disagree with Pruss (1998)’s suggestion that all circular explanations involved in consistent time travel scenarios fail to be self-explanatory. Some consistent time travel stories are self-explanatory, they fully explain and thus “determine” the items, and they are not hypodoxical. We might want to say involve “virtuous circles”. Cf. Billon (2019, especially fn.22) on virtuous circles and hypodoxes.

that the stick has been adjusted by another stick-adjuster who traveled through time from 2020 to 1999 with the l_1 -long stick and adjusted it. Here, as before, a stick-adjuster is someone who takes a stick of length x and

- (i) if $l \leq x \leq B$, cuts it by $\frac{x-l}{2}$ so that its size becomes $x - \frac{x-l}{2}$.
- (ii) if $b \leq x \leq l$, extends it by $\frac{l-x}{2}$ so that its size becomes $x + \frac{l-x}{2}$.

It takes little reflection to realize that given the behavior of the stick-adjusters the stick must be l -long both in 1999 and in 2020. Indeed, by construction $l_2 = l + 2(l_1 - l) = l + 4(l_2 - l)$ and likewise $l_1 = l + 2(l_1 - l) = l + 4(l_1 - l)$, which implies (I let the reader solve these two equations) that $l_1 = l = l_2$. The necessitation condition is satisfied which admittedly shows that the circular explanation is able not only to explain l_1 given l_2 and vice-versa, but l_1 and l_2 simpliciter. It is self-explanatory. The reader can develop a circular version of the Wheel-Turners infinite explanation.

Just like there are infinite explanations that are self-explanatory, then, there are circular explanations that are self-explanatory. At that point a foe of self-explanatory explanations might argue that our alleged examples of such self-explanatory explanations are all very peculiar, and that this peculiarity makes them somehow defective, illegitimate or at least uninteresting. They indeed share a striking formal feature. We can call explanatory laws, the laws that, in an infinite series of explanation, explain the transition from one item to the next. These laws are specified in the cases we have considered by the description of the behavior of scribes, the stick-adjusters or the wheel-turner. In ordinary explanations, the laws only explain a “final condition” *in combination with an “initial condition”*. In all the cases we have considered, however:

- *Historical Irrelevance*. Each item of the (infinite or circular) series is explained by the explanatory laws alone, rather than by the explanatory laws *and some initial condition*. (A connected epistemic feature: it is not necessary to know the precise history of the system (the successors of item #1) to figure out what the present state of the system is (item #1)).

It should be reminded that Historical Irrelevance does not entail that items in the series do not depend on their successors (in the sense that a change in the latter would not imply a change in the former). Historical irrelevance does not entail either that all items are constant. The first point was shown by the Stick-Adjusters and the Wheel-Turners examples, the second point is shown by the Sliding Monomaniacal Scribes and the Sliding Stick-Adjusters examples.

In any case, the objection goes, Historical Irrelevance is bad. In order to assess this criticism, we should ask two questions. We should ask, first, whether all infinite or circular self-explanatory explanations display Historical Irrelevance. We should then wonder whether all explanations that likewise display Historical Irrelevance are defective. I shall argue in

the next two sections that the answer to these two questions is respectively “yes” and “no”. All self-explanatory circular and infinite explanations display Historical Irrelevance, but this should not disqualify them as proper, legitimate explanations.

4 Self-explanations, Laws and Historical Irrelevance

I have tried to avoid committing as much as I could to a specific model of explanation so far. When I have put forward the necessitation condition, for example, I presented it as a contextual condition on explanations similar to that of the Regular Scribe example (the criterion, remember, relied on the claim that the explanation in that case failed to be self-explanatory because and only because what explained the transitions from one state to its antecedent failed to necessitate all the items of the series ; it said that an explanation of the same form would be self-explanatory if what explains the transitions managed to necessitate all the items of the series). I was also neutral on the kind of necessity involved.

In order to find out whether all self-explanatory circular or infinite explanations display Historical Irrelevance, let us now try to articulate general conditions on such explanations.

An explanation, I will assume, has always following form. It involves, first, an explanans constituted of:

- an initial condition
- and some explanatory laws L .

It involves an explanandum which is a just final condition. It involves, lastly, an explanatory relation \Rightarrow connecting the explanans to the explanandum. On the classical ND-models of explanation, for example, items are sentences or propositions and \Rightarrow is a priori deductibility or necessitation (strict implication).

A few precisions are in order. First the “initial” and “final” in “initial and final conditions” need not be read temporally. The priority here is explanatory, it will often be temporal as well when the explanation is a standard causal explanation, but it need not be.¹⁶

Second, the laws L need not be natural laws, they can be metaphysical “laws” or principles such as (arguably) the necessity of origins or the identity of indiscernibles, they can be logical “laws” or principles, mathematical “laws” or principles (axioms and definitions), moral laws, or just any laws you like, in fact (from now-on I will always prefer the word “laws” to the word “principle”). Laws need not be explicit either. In naïve, folk explanations the laws are often left implicit. This is probably the case of some metaphysical explanations as well. The fact that the laws are left implicit does not entail that they are totally unknown

¹⁶The notion of explanatory *priority* makes sense because explanation is supposed to yield a partial order.

or that they play no role in the explanation. There is in fact a long tradition in picturing metaphysical explanations as relying, just like naive mathematical explanations, on implicit principles that are known a priori.

Third, in some cases, the initial condition may be empty. For example, theorems in mathematics are explained by the axioms and definitions alone, and these can be considered as mathematical laws rather than as initial conditions.

Fourth there are different kinds of explanations, which depend on the kind of laws they rely on. An explanation will only be causal, for example, if the laws are natural laws (although, as we shall see, the converse might not be true). It will only be mathematical or metaphysical if the laws are mathematical or metaphysical, etc.

Now, consider an arbitrary infinite explanation: for all i , each item $\#i$ is explained by its successor $\#i + 1$ and the laws L_i . We can use “ U_i ” to abbreviate “the item $\#i$ ”. In Leibniz’s Regular Scribes example and our Monomaniacal Scribes example, U_i is the content of book copy $\#i$ and the explanatory laws $(L_i)_i$ are the laws describing the behavior of our scribes $(s_i)_i$. If our series is a series of successive states of the universe, U_i is the state of the universe at time t_i , and L_i are the cosmological laws between t_i and t_{i-1} (it is plausible that the cosmological laws remain constant but we do not need to assume that they do). Supposing that each state of the universe U_i is explained by cosmological laws L_{i+1} and its successor state U_{i+1} , we will also have an infinite explanation here, and arguably an infinite causal explanation. If our series is the series of successive grounds of the state of the Universe, L_i will be metaphysical laws that explain the transition from a ground $\#(i + 1)$ to what it grounds (the ground $\#i$ of the series if $i > 1$, or the state of the Universe if $i = 1$), etc.

By definition, an infinite explanation such as these will be self-explanatory iff by explaining the transitions from each item to its antecedent one explains the whole series, and so, iff what explains these transitions explain the whole series. But what explains these transitions are the laws $(L_i)_i$. So an infinite explanation will be self-explanatory iff

(C) All items are explained by the laws $(L_i)_i$ alone.¹⁷

Accordingly, all self-explanatory infinite explanations will indeed display Historical Irrelevance. The laws alone will suffice to explain the first item, and the history of the latter will be irrelevant.

A similar reasoning shows that self-explanatory circular explanations are those that satisfy (C), and that they display Historical irrelevance as well (just repeat the above reasoning

¹⁷This condition, it should be emphasized, does not entail that each law L_i can by itself explain one item. This last condition will only be satisfied in what we have called “degenerate self-explanatory explanations”, such as the Monomaniacal Scribes case.

supposing that the ranks of our items belonging to the ring of integers modulo n , $\mathbb{Z}/n\mathbb{Z}$, rather than to that of (non-null) natural numbers \mathbb{N}^*).

5 Infinite, Circular, Causal and Equilibrium Explanations

Let us admit then that all self-explanatory infinite or circular explanations display Historical irrelevance. Does that make them illegitimate? There are different ways to frame the accusation of illegitimacy. One could wonder, first, whether an explanation displaying Historical Irrelevance can really be a *causal explanations* and further claim that non-causal explanations are defective or illegitimate. After all, Historical Irrelevance suggests that even if an item $\#i$ may depend on its successor $\#i + 1$ in the sense that changing the latter would change the former, the nature of the first item does not really or heavily depend on that of its successors because it can be explained by the explanatory laws alone. This might cast a doubt on its being genuinely causal.

In answer, I should first emphasise that I have deliberately focused on explanations rather than causal explanations because I believe that non-causal explanations are ubiquitous and legitimate (there are explanations in mathematics, in metaphysics and in ethics that are unlikely to be causal) and I am not sure that we have a better grip of the notion of cause than on the broader notion of explanation.

Moreover, the cosmological argument I have put forward has, as we have seen, causal and non-causal readings. Yet, the non-causal readings seem more cogent to many (they seemed more cogent to Leibniz and Clarke, for example, and many theists and atheists have followed them on that point¹⁸). An advocate of cosmological arguments would thus probably be ill-advised to discount an explanation because it is not clearly causal.

Third, and more importantly, I think that even if it must be acknowledged that Historical Irrelevance will always make circular or infinite explanations that are self-explanatory peculiar, and set them apart from our many ordinary explanations, it is far from clear that such explanations cannot be causal, let alone legitimate. The fact that self-explanatory infinite or circular explanations are non-standard was to be expected: after all, we are not dealing, precisely, with our most ordinary explanations. Moreover, there are some perfectly kosher (if not perfectly ordinary) explanations that share the peculiarity of self-explanatory infinite and circular explanations, namely, *equilibrium explanations*. The statistician Ronald Fisher famously explained why the sex ratio of males and females is approximately one by the fact that any deviation from this ratio would be progressively cancelled by natural selection. This is a classical equilibrium explanation and it displays, like our self-explanatory

¹⁸Cameron (2008, 9) for examples affirms without further ado that, “If there is no first cause, if every causal chain is infinitely descending, we are not at all tempted to say that nothing has a causal explanation; but if there is no ultimate fundamentality, there is the strong intuition that nothing has a metaphysical explanation.”

infinite and circular explanations, a form of Historical Irrelevance: to the extent that this explanation is correct, the sex-ratio should always have been approximately one, and one can deduce that it is approximately one today without inquiring about its former values. Interestingly, while some philosophers believe that equilibrium explanations cannot be causal because of their Historical Irrelevance, others have argued some of them are good old causal explanations (see Strevens (2008, 267-72)).

There is in fact more than an analogy between equilibrium explanations and the kind of “self-explanation” involved in self-explanatory infinite or circular series. This self-explanation *is* an equilibrium explanation. Indeed, an equilibrium explanation is simply an explanation in which the final condition (the equilibrium) is explained by the explanatory laws of the system alone¹⁹ and is insensitive to the initial conditions. Such explanations are common in so-called “dynamical systems” (systems in which a function describes the time dependence of a point in a geometrical space). It is interesting to note that just like there are natural (and, as we have seen, arguably causal) equilibrium explanations, there might be non-natural, and for example metaphysical, equilibrium explanations. These are explanations in which the state of the system is explained by the metaphysical laws alone. The complete explanation why the world exists and why it is the way it is, I claim, might precisely be such an equilibrium explanation.

To sum up: even though they display a form Historical Irrelevance, self-explanatory infinite and circular explanations are, like equilibrium explanations, perfectly good explanations and the jury is still out on whether such explanations, when they involve natural, causal laws, are causal or not.

6 Explanation and Necessitation

We have argued that some infinite explanations are self-explanatory, we have defused an important objection from Alexander Pruss against this claim by showing that some circular explanations are self-explanatory as well. Finally, we have shown that all self-explanatory infinite or circular explanations will be peculiar in that they will display Historical Irrelevance, but that this should not disqualify them. Before moving on, I would like to come back to what might appear like a weak spot in our argument, namely the necessitation condition.

The real condition for a circular and infinite explanation be self-explanatory is, as we have seen, that it satisfies (C), i.e. that the explanatory laws alone *explain* (rather than necessitate) the whole series. As it seemed, however, that in Leibniz’s Regular Scribe example the infinite explanation only failed to be self-explanatory because the explanatory laws (or the behavior of the scribes they describe) failed to necessitate the whole series, we

¹⁹Or almost alone, when there is more than one stable equilibrium of the system.

have assumed that explanations of the same form would be self-explanatory provided that the explanatory laws *necessitate* (rather than explain) the whole series.

Now, there are classical objections against the claim that necessitation is always sufficient for explanation. Maybe these objections could add grist to the theist's mill and help him save the cosmological argument from the threat of self-explanatory infinite explanations. He could argue that in the infinite explanations of our examples, the explanatory laws necessitate all items but do not explain them. He could contend that what we have shown is at best that some infinite explanations might be "self-necessitating" (i.e. what explains all the transitions, makes the whole series necessary), not that they might be self-explanatory.

The relevant objections against the sufficiency of necessitation for explanation involve a form of explanatory overdetermination: when two independent items A and B both necessitate the same item C, at least one of them will normally be explanatory idle, and fail to explain C.²⁰ Imagine that John wants to kill Bob. In order to do that, he puts a bomb in the rear of Bob's car — call it Bomb1 — set to explode at the next morning at precisely 8:30 AM and 0 seconds. But John wants to be absolutely certain that Bob will die and when he thinks about it, during the night, he decides to put a second bomb — call it Bomb2 — in the front of Bob's car. Bomb2 is set to explode at precisely 8:30 AM and 1 second. Now suppose the state of Bob and Bomb1 at 8:00 and the explanatory laws make it necessary that John is dead at 8:31. Suppose that this is true as well of Bomb2: the explanatory laws and the state of Bob and Bomb 2 at 8:00 make it necessary that Bob will be dead at 8:31 AM. In such a case the antecedent state of Bomb2 and the explanatory laws would necessitate Bob's death but they would not explain it, because they are irrelevant. It is Bomb1 that explains the death. I suspect the theist could, likewise, invoke a form of explanatory overkill. He could argue that even if the infinite series involved in the cosmological argument were self-necessitating, i.e. even if what explains the transitions from one item to its antecedent makes the whole series necessary, we would still need to posit the existence of God to *explain* the whole series.

It is, however, extremely dubious that the friend of the cosmological argument could invoke infinite series that are self-necessitating but not self-explanatory to save her theistic conclusion. For once it is granted that an infinite explanation makes the whole series necessary, i.e. that the (non-godly) things that explain the transition from each item to its antecedent make the whole series necessary, the urge to posit a god to explain this whole series almost vanishes entirely. The fact that something A makes something C necessary might not imply that A explains C, but it implies that A is a good candidate for explaining C, and that, other things being equal, we do not need to posit some other entity to explain it.

²⁰There are other objections against the claim that necessitation suffices for explanation, objections based, for example, on explanatory asymmetries (see e.g. Woodward (2017)) but they do not seem to be relevant here.

7 Self-Explanatory Infinite vs. Theistic Explanations

Here is where we are. We have put forward some examples of self-explanatory infinite explanations. We have also argued that even though some infinite series might be self-necessitating without being self-explanatory, that should be of little help for the advocate of the cosmological argument. We have finally shown that while all self-explanatory infinite explanations display Historical Irrelevance, that should not disqualify them as legitimate explanations. At this point the Theist might opt for a more modest objection, and claim that even though self-explanatory infinite explanations are clearly possible, we cannot invoke them to defuse the theistic conclusion of the cosmological argument because they are always inferior to theistic explanations.

I see just three ways to articulate this objection. The first is based on the idea that infinite explanations require an unbounded time in the past but that (according to current cosmology) time is bounded in the past. The second relies on the claim that non-theistic infinite explanations must ultimately rely on unexplained laws. The third way is grounded on the claim claim that condition (C) is inconsistent with other aspects of our cosmological knowledge such as conservation principles.

Before tackling these objections let us note that the first and the third one only apply, by definition, to causal versions of the cosmological argument (and we have seen that there are other, non-causal and arguably more important versions of the argument). We should also emphasise that infinite explanations have at least one manifest advantage over theistic ones: unlike theistic explanations, they are, or rather, they can be completely naturalistic (they do not need to posit powers and beings too different in kind from those ascertained by the natural sciences).

7.1 Unbounded Past and Infinite (Natural) Explanations

Now, consider the first objection, based on the premises that an infinite explanation implies an unbounded time in the past and that this is ruled out by current cosmology. Both premises are false. The first premise seems to rely on a confusion between the unbounded and the infinite. An infinite (natural, causal) explanation does not imply an unbounded time in the past as the intervals $t_i - t_{i+1}$ between the state $\#i + 1$ and the state $\#i$ can decrease sufficiently quickly for their sum to be bounded. In fact, on standard Big-Bang models time is bounded but *open* in the past, so that there is no first moment, and any descending chain of explanation must be either infinite or fail to explain some (very) early events (Earman, 1995, 207-9). Moreover, it simply cannot be claimed that our current scientific knowledge implies that time is bounded in the past. As emphasised by Dowden (2019, §3) in a recent encyclopedia entry on the topic, the question of the amount of time in the past of the Universe is still debated by physicists, but there is a bit more support for the scenario which requires time to be unbounded.

7.2 Unexplained Laws

The Theist might also claim that inasmuch as they aim to remain naturalist, self-explanatory infinite explanations will not, like theistic explanations, be capable of providing a complete explanation. The idea, here, is that the best an infinite self-explanatory explanation can do is explain every state of the Universe from certain laws, and that these laws themselves call for an explanation. This criticism seems to rely on an unfair comparison, however. As noticed many times by Adolf Grünbaum (see e.g. his (Grünbaum, 1989)), theistic explanations are still extremely inchoate. In particular, they do not make explicit the details of the mechanisms by which God supposedly caused everything to be. One might likewise worry that once they are provided, these details will themselves call for an explanation and show that the theistic explanation is incomplete after all. Some Theistic explanations do provide, it is true, precise reasons grounding the existence of the Universe (if not precise causes of its existence). But one might still ask what grounds these reasons as well. Some Theists argue, for example, that reality is as good as it could be and that God wanted the best to be. But one might also ask why is that true? Is it because it is best that He wanted the best to be? If so, why is this latter claim true? etc.²¹

Conversely, non-theistic infinite explanations need not rely on laws that call for an explanation. To take the case of causal explanations, even though the natural laws we currently deem basic seem to call for an explanation, physicists are still hoping to find a “theory of everything” which would be, in the words of Einstein “rationally completely determined” (see Schilpp (1949, 63)), or even, as Hawking and Mlodinow (2010, 181) put it, be derived from “abstract considerations of logic alone”, and that would thus, arguably, not call for an explanation (Hawking and Mlodinow (2010) contend that they might well have found such a theory). Provided that we evaluate theistic and non-theistic explanations with the same standards, there is accordingly no reason to believe that non-theistic explanations are less likely to provide a complete explanation.

7.3 (C) and our Cosmological Knowledge

The last objection concerns, like the first one, the scientific plausibility of self-explanatory explanations. It contends that to our current cosmological knowledge is inconsistent with condition (C) and hence with the claim that an infinite explanation of the present state of Universe can be self-explanatory.

One could argue, for example, that present day cosmology appeals to past states (and not just to laws) in order to explain the present state of the universe. This is not totally true: as we have just seen, some cosmologists in search of a theory of everything, such as Hawking and Mlodinow (2010), do avoid or try to avoid any appeal to states. Others in the same

²¹Compare with the way Parfit (2011, vol. II, Appendix D) presses the so-called “axiarchic views” of the universe to explain why what is best should be.

research program believe that one day we will manage to avoid any appeal to states. In any case, the fact that many present day cosmological theories do appeal to states could only show that our present cosmological *knowledge* is inconsistent with (C) if these theories could reasonably be deemed:

- to provide us with *knowledge*
- and with knowledge of the history of the universe that is *scientifically complete* in the sense that it explains every cosmological fact that can be scientifically explained.

It seems, however, that inasmuch as they do pretend to be scientifically complete, present day cosmological theories (i) do not clearly provide *knowledge* (they rather provide interesting and plausible but speculative hypothesis) (ii) and do not, as we have emphasized, unanimously appeal to states (as opposed to laws only).

A more precise objection against the claim that our cosmological knowledge is consistent with (C) involves the conservation of energy. It says that the total energy of the universe being governed by a conservation law, if our series of items are series of energy levels of the universe at different times, our infinite explanation will, just like in the Regular Scribes case, explain why the energy of the Universe is the same all along, not why it has the value it does have. Accordingly, an infinite explanation of the energy level of the universe will not be self-explanatory. A fortiori, an infinite explanation of the state of the universe will not be self-explanatory.

One problem with this objection is that according to our best theories, it is not clear that the energy *of the whole universe* is conserved, nor, for that matter, that it is finite and well defined (see for example the excellent summary by Pitts (2019, §4) on the way conservation principles should be cautiously qualified).

A second problem with this objection is that the conservation principles, if they applied to the whole Universe would be likely to rule out theistic explanations as well : the creation of the universe by God, as it is traditionally understood (i.e. as being *ex nihilo*) seems indeed to violate conservation principles.

The morals here is that the question whether cosmological laws satisfy (C) is decisive but non-trivial, and, more broadly, that it is far from clear that we could come up with a scientific argument ruling out self-explanatory infinite explanations while invoking standards of scientific adequacy *that do not equally allow us to rule out the theistic hypothesis of a self-explanatory creator*.

8 Conclusion

Infinite explanations can provide simple, elegant and even naturalist alternatives to theistic explanations of reality. Were they not suspected of never being self-explanatory and thus of

always being incomplete, they would have many advocates today. Infinite explanations can, however, be self-explanatory, and when they are, they are as likely to eliminate lingering “why” questions as theistic explanations. They should be rehabilitated, and we should urgently consider how they can amend our causal and non-causal understanding of the cosmos.²²²³

References

- Billon, A. (2019). Paradoxical hypodoxes. *Synthese*, 196:5205–5229.
- Bliss, R. L. (2013). Viciousness and the structure of reality. *Philosophical Studies*, 166(2):399–418.
- Cameron, R. (2018). Infinite regress arguments. In *The Stanford Encyclopedia of Philosophy*.
- Cameron, R. P. (2008). Turtles all the way down: Regress, priority and fundamentality. *Philosophical Quarterly*, 58(230):1–14.
- Clarke, S. (1998). *A Demonstration of the Being and Attributes of God and Other Writings*. Cambridge University Press.
- Dasgupta, S. (2016). Metaphysical rationalism. *Noûs*, 50(2):379–418.

²²Even though I have concentrated on (some version of) the cosmological argument, here, the existence of self-explanatory infinite explanations has consequences that outreach philosophical cosmology. Cameron (2008) has for example put forward an elegant argument for the claim that infinite descending chains of explanations such as the one that we have considered involve a *vicious* regress.

It would be better to be able to give a common metaphysical explanation for every dependent entity [every item in the chain that is explained by another one]. We can do that only if every dependent entity has its ultimate onto-logical basis in some collection of independent entities; so this provides reason to believe the intuition against infinite descent in metaphysical explanation (Cameron, 2008, 12).

(...) what needs to be explained is the existence of each dependent entity; and while (...) in an infinitely descending chain of ontologically dependent entities, there is an explanation for why each dependent entity exists, there is no single explanation for why all the dependent entities exist (Cameron, 2018).

According to him an infinite explanation would always lack unity (and simplicity) because it would fail to provide a common explanation of each of the item. As his argument is based on theory-choice considerations, Cameron does not conclude that such infinite explanations are impossible, but only that they do not occur in our world. As I understand it, however, his argument crucially depends on the claim that an infinite explanation cannot be self-explanatory. Indeed, in our examples of self-explanatory infinite explanations, in virtue of the fact that each item of the series is explained by its successor (and the explanatory laws), there *is* a common explanation of all items of the series (by the explanatory laws alone).

²³This paper has been presented at the Institut Jean Nicod Colloquium in June 2020. I am also indebted to Alexander Pruss and Ghislain Guignon for their comments on an earlier draft.

- Dowden, B. (2019). What else science requires of time. In Fieser, J. and Dowden, B., editors, *The Internet Encyclopedia of Philosophy*.
- Earman, J. (1995). *Bangs, crunches, whimpers, and shrieks: Singularities and acausalities in relativistic spacetimes*. Oxford University Press.
- Eldridge-Smith, P. (2007). Paradoxes and hypodoxes of time travel. In Jones, J. L., Campbell, P., and Wylie, P., editors, *Art and Time*, pages 172–189. Australian Scholarly Publishing.
- Fine, K. (2012). Guide to ground. In Correia, F. and Schnieder, B., editors, *Metaphysical Grounding*, pages 37–80. Cambridge University Press.
- Grünbaum, A. (1989). The pseudo-problem of creation in physical cosmology. *Epistemologia*, 12(1):3.
- Guigon, G. (2015). A universe of explanations. In Bennett, K. and Zimmerman, D. W., editors, *Oxford Studies in Metaphysics*, pages 345–375. Oxford University Press.
- Hawking, S. and Mlodinow, L. (2010). *The Grand Design*. Bantam Books.
- Hudson, H. (1997). Brute facts. *Australasian Journal of Philosophy*, 75(1):77–82.
- Hume, D. (1907). *Dialogues concerning natural religion*. William Blackwood.
- Leibniz, G. W. (1989). On the radical origination of things. In *Philosophical Papers and Letters*, pages 486–491. Springer.
- Oppy, G. (2006). *Philosophical Perspectives on Infinity*. Cambridge University Press.
- Parfit, D. (2011). *On what matters*. Oxford Univ Press, New York.
- Pitts, J. B. (2019). Conservation laws and the philosophy of mind: Opening the black box, finding a mirror. *Philosophia*.
- Pruss, A. R. (1998). The hume-edwards principle and the cosmological argument. *International Journal for Philosophy of Religion*, 43(3):149–165.
- Pruss, A. R. (2006). *The principle of sufficient reason: A reassessment*. Cambridge University Press.
- Pruss, A. R. (2018a). Causal finitism and the kalaam argument.
- Pruss, A. R. (2018b). *Infinity, Causation, and Paradox*. Oxford University Press.
- Robinson, A. (1961). *Non-Standard Analysis*. North-Holland Publishing Co.

- Rowe, W. L. (1970). Two criticisms of the cosmological argument. *The Monist*, 54(3):441–459.
- Rowe, W. L. (2001). *Philosophy of Religion: An Introduction*. Wadsworth/Thomson Learning.
- Russell, B. (1927). *Why I Am Not a Christian*. Routledge.
- Schilpp, P. A. (1949). *The Library of Living Philosophers, Volume 7. Albert Einstein: Philosopher-Scientist*. Open Court.
- Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press.
- Swinburne, R. (2004). *The existence of God*. Oxford University Press.
- Woodward, J. (2017). Scientific explanation. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2017 edition.